REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
24-06-2013	Final Report	1-Sep-2012 - 31-May-2013

4. TITLE AND SUBTITLE
Biomolecular Principles of Matrix Assembly Related to Fracture Resistance

5a. CONTRACT NUMBER
W911NF-12-1-0255

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
John Spencer Evans

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
New York University
Office of Sponsored Programs
15 Washington Place, 1-H
New York, NY 10003 -6657

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
62175-LS-II.1

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
One of the major engineering feats of the developing spicule matrix is the assembly of a protein scaffolding network that readily adapts itself to the emergence of ACC clusters and eventually persists within an intracrystalline environment as the ACC phase transforms into crystalline calcite. Collectively, our findings indicate that SM50 is suitably adapted for a major role in this process. We confirm that rSM50 spontaneously oligomerizes to form amorphous, heterogeneous supramolecular protein complexes that can form films and behave

15. SUBJECT TERMS
self-assembly, spicule matrix, sea urchin, fracture resistance, amorphous materials

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			John Evans
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 212-998-9605

Report Title

Biomolecular Principles of Matrix Assembly Related to Fracture Resistance

ABSTRACT

One of the major engineering feats of the developing spicule matrix is the assembly of a protein scaffolding network that readily adapts itself to the emergence of ACC clusters and eventually persists within an intracrystalline environment as the ACC phase transforms into crystalline calcite. Collectively, our findings indicate that SM50 is suitably adapted for a major role in this process. We confirm that rSM50 spontaneously oligomerizes to form amorphous, heterogeneous supramolecular protein complexes that can form films and behave in a relatively mobile fashion. This would provide a means for quickly assembling a protein matrix with fluid or labile features that are commensurate with those of the ACC phase itself. Moreover, the lability of rSM50 assemblies would provide an adaptation to the changing shape and dimension of spicules as they undergo developmental elongation and maturation, i.e., the SM50-dominated spicule matrix would be "plastic" for all intents and purposes and thus is perfectly suited for embryonic development and eventual mineralization, with the added benefit of providing a cushioning or compressive phase as fracture-resistant intracrystalline components within crystalline calcite.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Poster presentation, GRC Thins Films and Crystal Growth Conference, July 2013

Number of Presentations: 1.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received Paper

TOTAL:

Number of Manuscripts:

Books

Received Paper

TOTAL:

Patents Submitted

None

Patents Awarded

None

Awards

none

Graduate Students

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Iva Perovic	0.50
FTE Equivalent:	**0.50**
Total Number:	**1**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
John S. Evans	0.25	
FTE Equivalent:	**0.25**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Michael Lui	0.10	Chemistry
Joseph Wu	0.10	Chemistry
FTE Equivalent:	**0.20**	
Total Number:	**2**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	0.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....	2.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....	0.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	2.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME PERCENT_SUPPORTED

FTE Equivalent:

Total Number:

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Technology Transfer

Final Report, ARO Short Term Innovation Research (STIR)

Period: 09/01/12 – 5/31/13
PI: John Spencer Evans
Institution: New York University
Grant Award: W911NF-12-1-0255

Background: The sea urchin represents an important research model for understanding a wide variety of phenomena, including molecular biology, evolution, and biomaterial formation. With regard to the latter, the species *Strongylocentrotus purpuratus* (purple sea urchin) and its calcium carbonate-containing adult spine and embryonic spicule skeletal elements have provided insights into the formation, stabilization, and crystalline transformation of amorphous minerals.[1-10] This mineralization process is highly interesting and in the sea urchin spicule it is now believed that hydrated and dehydrated variants of amorphous calcium carbonate (ACC) are the true precursors to crystalline calcite.[1] Even more interesting is the fact that in certain regions of the spicule the hydrated ACC variant persists as stabilized nanoparticles alongside calcite crystals.[1] Recent studies have shown that the spicule matrix (SM) proteins of *S. purpuratus* embryonic spicules are physically associated with ACC deposits and may be involved in ACC localization and its eventual transformation.[1-10] Interestingly, the entrapment of these proteins within the mineral phase is one of the major contributors to the fracture resistance of these skeletal elements.[6-9] Hence, there is a great interest in understanding the functional role(s) of these proteins within the spiculogenesis process, not only from a biological standpoint, but also from a materials perspective that is relevant to the mission of the US Army Research Office.

With the successful sequencing of the *S. purpuratus* genome,[11-13] we now know that there are 16 unique SM biomineral-associated genes. Via RNA splicing pathways, these genes code for > 40 expressed matrix proteins[11-14] that are secreted by the primary mesenchyme cells into a membrane-bound mineralization space where the spicule will form.[5,8,12,13] All expressed SM proteins feature a canonical structure consisting of a N-terminal leader sequence and a C-type lectin-like domain (CTLL),[11-14] and in 10 SM proteins there also exists a C-terminal Gln, Pro, Gly-rich repetitive domain.[11-14] It is known that the SM proteins assemble to form a concentric

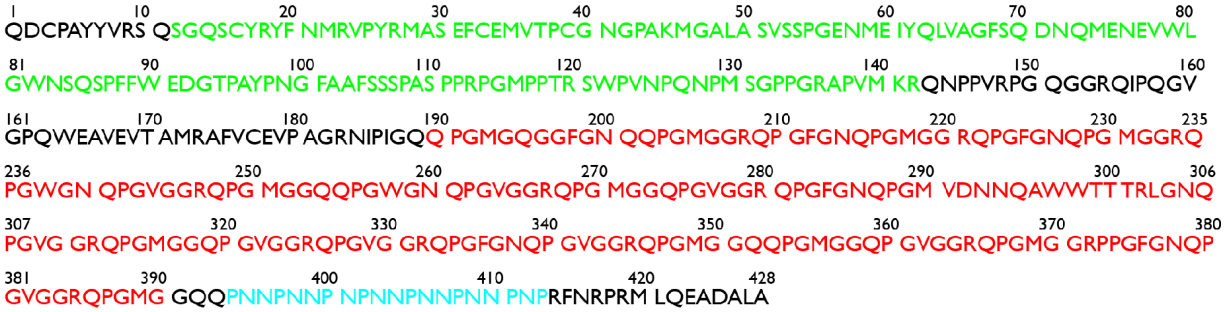


Figure 1: Primary sequence of mature S. *purpuratus* spicule matrix protein SM50. Color coding of sequence regions: Green = CTLL domain; Red = Gln, Pro, Gly-rich repeat region; Blue = Pro, Asn-rich region.

matrix that courses throughout the mineral phase of the spicule, with some SM proteins localized at the growing mineralization front and others participating in the secretion of spicule

components.[9,14-16] Thus, there are specific functional and regional roles for each SM protein and there is no doubt that spicule matrix assembly is a critical event in the mineralization process. However, very little is known regarding the spicule matrix protein assembly process and its organization, or, the involvement of SM proteins with ACC deposits.

To derive a better understanding of the overall spicule matrix assembly process, one must first understand the assembly behavior of individual SM proteins and then progress to mixed SM systems where heterogeneous protein – protein interactions can be evaluated. At present, the best candidate for individual study is SM50, a 44.5 kDa basic (pI = 10.73) single polypeptide

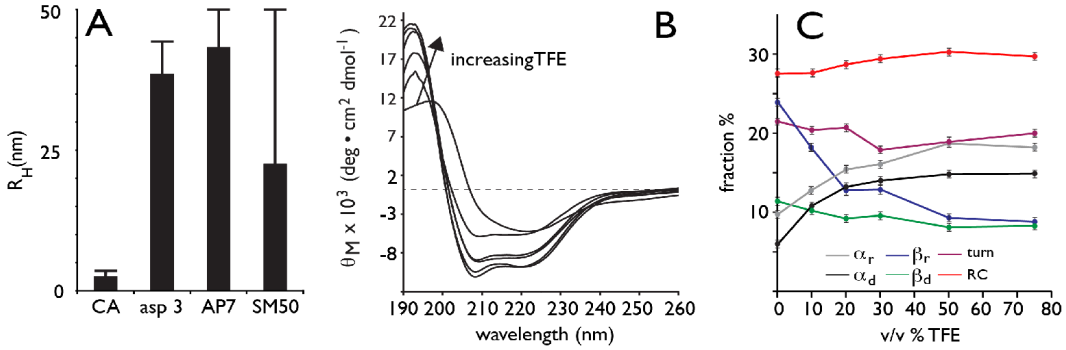
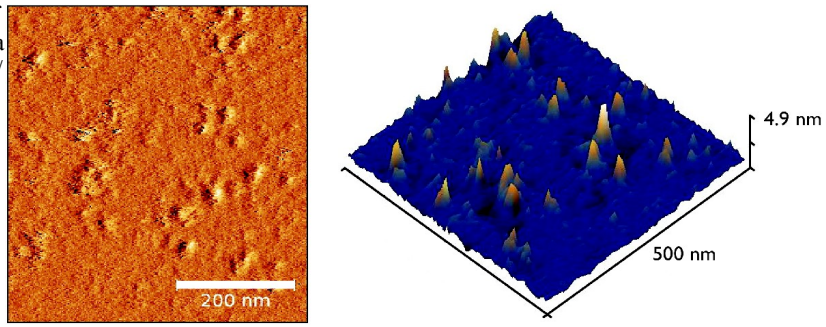


Figure 2: (A) DLS-determined hydrodynamic radii (R_H) for 164 µM apo-SM50 oligomers at pH 9.76 in 10 mM $NaHCO_3$ / Na_2CO_3 buffer. For comparison, we present R_H values obtained in the same buffer/pH solutions for bovine erythrocyte carbonic anhydrase II (Sigma/Aldrich, USA), A. *rigida* mollusk shell prismatic protein Asprich "3", and H. *rufescens* mollusk shell nacre protein, AP7. (B) Far UV circular dichroism spectrum of 6 µM SM50, pH 9.76 in 10 mM $NaHCO_3$ / Na_2CO_3 buffer, as a function of 2,2,2-trifluoroethanol (TFE) content (0, 10, 20 30, 50, 75% v/v). (C) Fraction % of secondary structures calculated by TFE titration of 6 µM SM50, pH 9.76 in 10 mM $NaHCO_3$ / Na_2CO_3 buffer.

(**Figure 1**) that is the most abundant protein in the embryonic spicule, the mature adult spine,[1,9,12] and the tooth and test skeletal elements of this sea urchin.[12-16] This protein is a member of a subfamily that includes SM37, SM32, SM29, PM27 and three predicted SM29-related spicule proteins.[11-14] In the spicule, SM50 is preferentially localized along the interior of the spicule sheath at the periphery of the mineral phase,[1,6,9,15,16] and thus it is believed to play a major role in ACC stabilization and transformation.[1,20-22] The primary structure of SM50 features the canonical CTLL domain within the N-terminal portion of the protein.[11-14] Intriguingly, the C-terminal domain contains two repetitive domains and a charged C-terminus. The first is a 203 AA Gln, Pro, Gly-rich consensus repeat sequence, –QPG(F/M/W)G(N/G)QPG(V/M)GG(R/Q)–,[12,14,17]

Figure 3: Representative AFM images of rSM50 oligomers (6 µM) forming on mica substrates at pH 9.76 in 10 mM $NaHCO_3$ / Na_2CO_3 buffer.



with the most common variants -GVGGR- and -GMGGQ- homologous to both elastin and spider dragline silk protein elastomeric repeats.[18,19] The second is a conformationally labile 20 AA Pro, Asn-rich repeat[19] that is upstream of the Gln, Pro, Gly-rich repeat.[12,14,17-19] Unfortunately, we know very little about the function(s) of the CTLL and repetitive domains, the

aggregation or assembly properties of SM50, or how protein self-assembly might facilitate ACC formation and stabilization.

Research: To resolve this, we initiated in vitro studies of apo-SM50 self-assembly using a recombinant form of SM50 (rSM50) and solution conditions approximating those of in vitro pre-nucleation cluster mineralization assays.[20-22] Under these conditions we find that apo-rSM50 is an intrinsically disordered protein (**Figure 2**) that is fold-inducible and assembles to form disordered supramolecular complexes that possess a high degree of dimensional heterogeneity (**Figures 2A, 3**). These protein assemblies are "plastic", i.e., they are highly dynamic with evidence of backbone and sidechain motion emanating from the repetitive Gln, Pro, Gly and Pro, Asn repeat domains of the C-terminal region (**Figure 4**). Interestingly, the N-terminal CTLL region does not exhibit this phenomenon. We note that dynamic, labile behavior is also common to other mineral-stabilizing biomineralization proteins assemblies[23-29] and to disordered polymer-induced liquid precursor (PILP) phases that stabilize amorphous minerals in vitro and regulate their transformation into crystalline solids.[30,31] Using bioinformatics, we confirm that the C-terminal Gln, Pro, Gly repetitive domain is the primary source of intrinsic disorder[23,32-38] in the SM50 sequence (**Figure 5**). Interestingly, bioinformatics predictions indicate that the N-terminal CTLL region of SM50 possesses a significant level of amyloid-like cross-beta strand regions (Figure 6),[23,39-43] which are important for protein-protein assembly and this implicates the CTLL domain in SM50 supramolecular assembly. Thus, SM50 is a disordered, aggregation-prone protein that forms

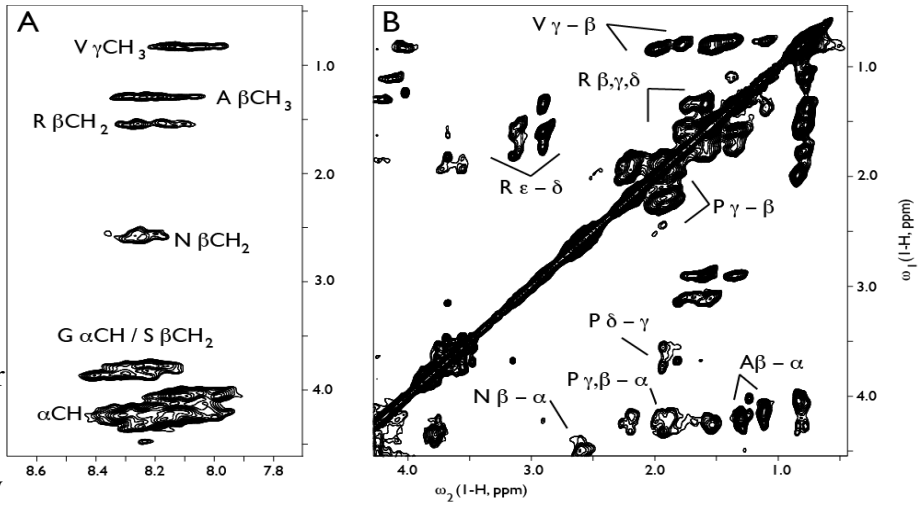


Figure 4: 600 MHz 1-H NMR TOCSY spectra of 116 μM rSM50 in UDDW, pH 7.5, (A) NH – CHα fingerprint region; (B) Sidechain aliphatic region. Tentative assignments of amino acid spins systems are indicated on the plots.

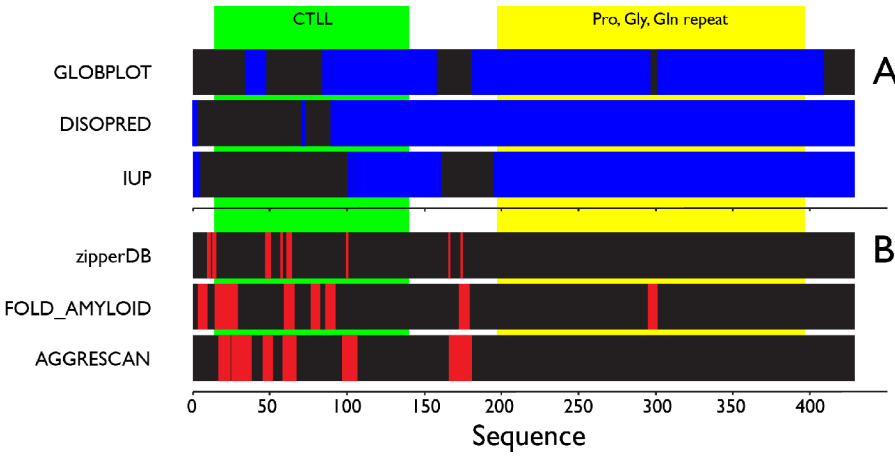


Figure 5: Graphical representations of predicted locations of (A) intrinsic disorder (GLOBPLOT, DISOPRED, IUP algorithms, blue color) and (B) amyloid-like aggregation prone (zipperDB, FOLD-AMYLOID, AGGRESCAN algorithms red color) regions of the SM50 mature sequence. The sequence locations of the CTLL (green) and Gln, Pro, Gly repetitive (yellow) domains are shown as overlays.

highly labile, dynamic PILP-like protein assemblies and these features may facilitate ACC localization in the embryonic spicule.

Conclusions: One of the major engineering feats of the developing spicule matrix is the assembly of a protein scaffolding network that readily adapts itself to the emergence of ACC clusters and eventually persists within an intracrystalline environment as the ACC phase transforms into crystalline calcite. Collectively, our findings indicate that SM50 is suitably adapted for a major role in this process. We confirm that rSM50 spontaneously oligomerizes to form amorphous, heterogeneous supramolecular protein complexes that can form films and behave in a relatively mobile fashion. This would provide a means for quickly assembling a protein matrix with fluid or labile features that are commensurate with those of the ACC phase itself. Moreover, the lability of rSM50 assemblies would provide an adaptation to the changing shape and dimension of spicules as they undergo developmental elongation and maturation, i.e., the SM50-dominated spicule matrix would be "plastic" for all intents and purposes and thus is perfectly suited for embryonic development and eventual mineralization, with the added benefit of providing a cushioning or compressive phase as fracture-resistant intracrystalline components within crystalline calcite.

Relevance of this research to ARO: Our findings provide a new blueprint for designing new composite materials that incorporate protein-inspired philosophies. First, intrinsic disorder is a primary component for designing "reactive" polymers that can form assemblies that will interact with other components, such as inorganic solids, at a later period in the composite assembly process. Second, the incorporation of amyloid-like sequence motifs can create specific docking sites for protein – protein or polymer-polymer interactions that can stabilize large supramolecular assemblies. The degree to which these two features are incorporated into future polymers or proteins could be used to "tune" the assembly process and the molecular features of the resultant assemblies.

Publications generated from this grant:

Iva Perovic, Joseph Z. Wu, Trinanjana Mandal, Michael Liu, Jong Seto, Helmut Cölfen, and John Spencer Evans (2013) A C-type lectin-like sea urchin biomineralization protein, SM50, oligomerizes to form dynamic film-like assemblies. *Biochemistry*, submitted.

References

1. Gong, Y.U.T., *et al*., (2012) *Proc. Natl. Acad. Sci. USA* **109**, 6088-6093.

2. Seto, J., *et al*., (2012) *Proc. Natl. Acad. Sci. USA* **109**, 3699-3704.

3. Wilt, F.H. (1999) *J. Struct. Biol.* **126**, 216-226.

4. Raz, S., *et al*., (2003) *Adv. Materials* **13**, 480-486.

5. Beniash, E., *et al.*, (1997) *Proc. R. Soc. Lond. B Biol.Sci.,* **264**, 461–465.
6. Aizenberg, J., *et al*., (1997) *J. Am. Chem. Soc.* **119**, 881-886.

7. Albeck, A., *et al.*, (1993) *J. Am. Chem. Soc.* **115**, 11691-11697.

8. Cho, J.W., Parin, J., Lennarz, W.J. (1996) *Proc. Natl. Acad. Sci USA* **93**, 1282-1286.

9. Seto, J., *et al.*, (2004) *J. Struct. Biol.* **148**, 123–130.

10. Politi Y., *et al*. (2008) *Proc. Natl. Acad. Sci. USA* **105**, 17362–17366.

11. Sodergren, E., *et al*., (2007) *Science* **314**, 941-952.

12. Mann, K., Wilt, F.H., Poustka, A.J. (2010) *Proteome Science* **8**, 33-35.

13. Mann, K., Poustka, A.J., Mann, M. (2008) *Proteome Science* **6**, 22-32.

14. Livingston, B.T., *et al*., (2006) *Devel. Biol.* **300**, 335-348.

15. Urry, L.A., *et al*., (2000) *Devel. Biol.* **225**, 201-213.

16. Ingersoll, E.P., McDonald, K.L., Wilt, F.H. (2003) *J. Exptl. Zoology* **300A**, 101-112.

17. Katoh-Fukui, Y., *et al*., (1991) *Devel. Biol.* **145**, 201-202.

18. Xu, G., Evans, J.S. (1999) *Biopolymers* **49**, 303-312.

19. Zhang, B., Xu, G., Evans, J.S. (2000) *Biopolymers* **54**, 464-475.

20. Gebauer, D., Volkel, A., Coelfen, H. (2008) *Science* **322**, 1819-1822.
21. Gebauer, D., Coelfen, H. (2011) *Nano Today* **6**, 564-584.
22. Gebauer, D., *et al*., (2009) *Advanced Materials* **21**, 435 – 439.
23. Evans, J.S. (2012) *Bioinformatics* **28**, 3182-3185.
24. Ndao, M., *et al*., (2011) *Protein Science* **20**, 724-734.
25. Amos, F.F, Ponce, C.B., Evans, J.S. (2011) *Biomacromolecules* **12**, 1883-1889.
26. Ponce, C.B., Evans, J.S. (2011) *Crystal Growth and Design* **11**, 4690-4696.
27. Amos, F.F., Evans, J.S. (2009) *Biochemistry* **48**, 1332-1339.
28. Amos, F.F., *et al*., (2011) *Biochemistry* **50**, 8880-8887.
29. Ndao, M., Ponce, C.B., Evans, J.S. (2012) *Faraday Discussions* **159**, 449-462.

30. Gower, L. B. (2008) *Chem. Rev.* **108**, 4551–4627.

31. Wolf, E., *et al*., (2011) *J. Am. Chem. Soc.* **133**, 12642-12649.

32. Peysselon, F., *et al*., (2011) *Mol. BioSyst.*, **7**, 3353-3365.

33. Linding, R., *et al*., (2003) *Nucleic Acids Res*. **31**, 3701-3708.

34. Dosztányi, Z., *et al*., (2005) *Bioinformatics* **21**, 3433-3434.

35. Ward, J.J., *et al*., (2004) *J. Mol. Biol.* **337**, 635-645.

36. Tompa, P., (2002) *Trends Biochem. Sci.* **27**, 527-533.

37. Uversky, V.N., Dunker, A.K. (2010) *Biochim. Biophys. Acta,* **1804**, 1231–1264.

38. Xie, H., *et al*., (2007) *J. Proteome Res.* **6**, 1882-1898.

39. Conchillo-Sole, O., *et al*., (2007) *BMC Bioinformatics* **8**, 65-82.

40. Garbuzynsky, S.O., *et al*., (2010) *Bioinformatics* **26**, 326-332.

41. Greenwald, J., Riek, J. (2012) *J. Mol. Biol.* **421**, 417-426.

42. Linding, R., *et al*., (2004) *J. Mol. Biol.* **342**, 345-353.

43. Bryan, A.W. *et al.*, (2009) *PLOS Comput. Biol.* **5**, 1-11.